OPEN

Emerging Microbes and Infections (2015) 4, e6; doi:10.1038/emi.2015.6

www.nature.com/emi

ORIGINAL ARTICLE

Identification of specific metabolites in culture supernatant of *Mycobacterium tuberculosis* using metabolomics: exploration of potential biomarkers

Susanna KP Lau[1,2,3,4], Ching-Wan Lam[5], Shirly OT Curreem[4], Kim-Chung Lee[4], Candy CY Lau[4], Wang-Ngai Chow[4], Antonio HY Ngan[4], Kelvin KW To[1,2,3,4], Jasper FW Chan[1,2,3,4], Ivan FN Hung[1,2,3,4], Wing-Cheong Yam[4], Kwok-Yung Yuen[1,2,3,4] and Patrick CY Woo[1,2,3,4]

Although previous studies have reported the use of metabolomics for *Mycobacterium* species differentiation, little is known about the potential of extracellular metabolites of *Mycobacterium tuberculosis* (MTB) as specific biomarkers. Using an optimized ultrahigh performance liquid chromatography–electrospray ionization–quadruple time of flight–mass spectrometry (UHPLC–ESI–Q–TOF–MS) platform, we characterized the extracellular metabolomes of culture supernatant of nine MTB strains and nine non-tuberculous *Mycobacterium* (NTM) strains (four *M. avium* complex, one *M. bovis* Bacillus Calmette–Guérin (BCG), one *M. chelonae*, one *M. fortuitum* and two *M. kansasii*). Principal component analysis readily distinguished the metabolomes between MTB and NTM. Using multivariate and univariate analysis, 24 metabolites with significantly higher levels in MTB were identified. While seven metabolites were identified by tandem mass spectrometry (MS/MS), the other 17 metabolites were unidentified by MS/MS against database matching, suggesting that they may be potentially novel compounds. One metabolite was identified as dexpanthenol, the alcohol analog of pantothenic acid (vitamin B5), which was not known to be produced by bacteria previously. Four metabolites were identified as 1-tuberculosinyladenosine (1-TbAd), a product of the virulence-associated enzyme Rv3378c, and three previously undescribed derivatives of 1-TbAd. Two derivatives differ from 1-TbAd by the ribose group of the nucleoside while the other likely differs by the base. The remaining two metabolites were identified as a tetrapeptide, Val-His-Glu-His, and a monoacylglycerophosphoglycerol, phosphatidylglycerol (PG) (16 : 0/0 : 0), respectively. Further studies on the chemical structure and biosynthetic pathway of these MTB-specific metabolites would help understand their biological functions. Studies on clinical samples from tuberculosis patients are required to explore for their potential role as diagnostic biomarkers.
Emerging Microbes and Infections (2015) **4,** e6; doi:10.1038/emi.2015.6; published online 28 January 2015

Keywords: biomarkers; diagnosis; metabolomics; *Mycobacterium tuberculosis*; specific

INTRODUCTION

Despite being an ancient disease, tuberculosis, caused by *Mycobacterium tuberculosis* (MTB), is still causing 9.0 million new cases and 1.5 million deaths worldwide in 2013, with one-third of the world's population thought to have been infected.[1] The disease is more prevalent in the developing world, especially in Asian and African countries. In Hong Kong, despite being a developed city, more than 4000 new cases are being reported per year. Since the 1980s, there has been re-emergence of tuberculosis in developed countries that coincided with the acquired immunodeficiency syndrome epidemic and later increasing use of immunosuppressants. Disease patterns have also changed, with a higher incidence of disseminated and extrapulmonary disease, coupled with the emergence of multidrug-resistant strains.[2–4] While MTB is the most common *Mycobacterium* species associated with infections in immunocompetent hosts, immunocompromised patients such as those with acquired immunodeficiency syndrome or type 1 cytokine pathway defects and transplant recipients are also at risk of infections by non-tuberculous *Mycobacterium* (NTM) species.[5–7]

The current gold standard for diagnosis of tuberculosis is smear and culture to look for acid-fast bacilli in clinical specimens. Although culture using solid or liquid medium is considered more sensitive and specific than smear, it is associated with significant pitfalls. First, culture-negative cases are often encountered when bacterial loads are low, especially in patients with early, disseminated or extrapulmonary disease, young children and immunocompromised patients.[8,9] Second, it usually takes at least 2–6 weeks before bacterial growth can be detected and even longer for species identification by phenotypic tests, which may result in treatment delay in smear-negative cases. As a result, novel diagnostic modalities such as adenosine deaminase levels in pleural fluid and nucleic acid amplification by polymerase chain reaction (PCR) have been developed to aid diagnosis. However, these methods are still far from optimal. For example, PCR has a reported sensitivity of only 60%–80% using culture

[1]State Key Laboratory of Emerging Infectious Diseases, Hong Kong, China; [2]Research Centre of Infection and Immunology, The University of Hong Kong, Hong Kong, China; [3]Carol Yu Centre for Infection, The University of Hong Kong, Hong Kong, China; [4]Department of Microbiology, The University of Hong Kong, Hong Kong, China and [5]Department of Pathology, The University of Hong Kong, Hong Kong, China
Correspondence: SKP Lau; PCY Woo
E-mail: skplau@hkucc.hku.hk; pcywoo@hkucc.hku.hk
Received 26 September 2014; revised 27 November 2014; accepted 26 December 2014

as the gold standard, and is often limited by the presence of PCR inhibitors in clinical specimens.[10,11] Therefore, the availability of alternative techniques for improved diagnosis of tuberculosis is eagerly awaited, and such techniques should be able to differentiate between MTB and NTM infections which necessitate different treatment regimens.

Metabolomics is an uprising research platform for systematic studies of the small-molecular metabolite profiles of a cell, tissue or organism, which are the end products of cellular processes. Using statistical analyses, the metabolic profiles from different cells or systems can be compared, which can be used to differentiate between different biological systems and identify potential metabolite markers specific to these systems. The technique has been applied to characterize various diseases or pathogens including MTB.[12–18] Using this approach, metabolomic data obtained from urine samples have also been used to distinguish healthy subjects from patients with infections such as pneumococcal disease and urinary tract infections.[19–21] However, previous metabolomics studies on MTB isolates were mainly focused on detection from culture and species/strain identification by analyses of intracellular metabolites.[13,14,22] Little is known about the potential of extracellular metabolites of MTB as specific biomarkers. For example, metabolomics studies have been performed to identify various *Mycobacterium* species, compare hyper- and hypo-virulent strains and study carbon utilization of MTB strains.[13,14,23] Although a few studies using samples from infected patients or animals have revealed potential signature metabolites, they are not yet confirmed to be useful routine diagnostic purposes.[24–26] Since MTB is able to produce volatile organic compounds and stearic acid which can be detected in the urine and sputum of infected patients respectively,[24,27] we hypothesize that there are potentially novel extracellular metabolites that are specifically produced by MTB that may be detected in body fluids for non-invasive diagnosis of tuberculosis. To search for potential biomarkers for diagnosis of tuberculosis, we attempted to characterize the metabolomes of culture supernatants of MTB and NTM species, using ultrahigh performance liquid chromatography–electrospray ionization–quadruple time of flight–mass spectrometry (UHPLC–ESI–Q–TOF–MS). Multi- and univariate statistical analyses of the metabolome data were used to identify specific metabolites that are found only in the culture supernatants of MTB but not NTM species.

MATERIAL AND METHODS

Mycobacterial strains and culture

Nine MTB and nine NTM (four *M. avium* complex strains, one *M. bovis* Bacillus Calmette–Guérin (BCG) strain, one *M. chelonae* strain, one *M. fortuitum* strain and two *M. kansasii* strains) strains were included in this study (Supplementary Table S1). All clinical isolates were identified by standard conventional methods.[28] Each *Mycobacterium* strain was grown on Lowenstein–Jensen solid medium for 3–4 weeks with continuous aeration. Colonies were scraped from the Lowenstein–Jensen slants and incubated in 30 mL Middlebrook 7H9 medium supplemented with 0.2% (v/v) glycerol and 10% oleic acid-albumin-dextrose-catalase (Becton-Dickinson, Sparks, MD, USA) without Tween-80 in filtered screw-cap Erlenmeyer culture flasks with baffled bottom (Thermo Scientific, Waltham, MA, USA) at 37 ℃ with shaking at 100 rpm for 2 days for *M. chelonae* and *M. fortuitum* and 21 days for all strains of MTB, *M. avium* complex, *M. bovis* and *M. kansasii*. The primary subcultures, adjusted to OD600 0.1, were further subcultured in 30 mL of fresh supplemented Middlebrook 7H9 liquid medium at 37 ℃ with shaking at 100 rpm for 4 days for *M. chelonae* and *M. fortuitum*; and 14 days for MTB, *M. avium* complex, *M. bovis* and *M. kansasii* until OD600 1.5 for all strains to achieve comparable bacterial concentrations. OD600 1.5 was chosen for harvest because it represents early stationary phase in MTB strains as shown in previous studies.[29] The secondary subcultures were centrifuged at 3000 rpm for 30 min to obtain the supernatant which was filtered twice using 0.22 μm filters (Millipore, Billerica, MA, USA). Metabolic activities in the filtrates were quenched immediately by incubating the filtrates in liquid nitrogen for 10 min. The filtrates were lyophilized and stored at −80 ℃ until sample extraction and analysis. Uninoculated culture medium was used as negative control. Three biological replicates of the cultures were used.

Chemicals and reagents

Liquid chromatography–mass spectrometry (LC–MS) grade water, methanol and acetonitrile were purchased from J.T. Baker (Center Valley, PA, USA). Analytical grade acetic acid, 5 M ammonium acetate and standard chemical dexpanthenol were purchased from Sigma-Aldrich, Inc. (St Louis, MO, USA),

Sample preparation

Lyophilized samples were reconstituted by dissolving in 1 mL solvent mixture containing water/methanol/acetonitrile (1 : 2 : 2). The samples were vortexed for 1 min and subsequently sonicated for 10 min at room temperature. After centrifugation at 15 000*g* for 15 min at 4 ℃, supernatants were transferred to LC vial for LC–MS analysis.

UHPLC–ESI–Q–TOF–MS

For LC, the separation was performed by Agilent 1290 UHPLC (Agilent Technologies, Santa Clara, CA, USA) and Agilent Eclipse Plus RRHD C18 (2.1×100 mm, 1.8 μm) column with Agilent Eclipse Plus RRHT C18 (2.1×30 mm, 1.8 μm) guard column. The injection volume was 3 μL. The column and autosampler temperature were maintained at 45 ℃ and 10 ℃, respectively. The separation was performed at a flow rate of 0.4 mL/min under a gradient program in which mobile phase A was composed of 5 mM ammonium acetate in water containing 0.1% acetic acid (v/v) and mobile phase B was composed of 0.05% acetic acid (v/v) in acetonitrile. The gradient program was applied as follows: t=0 min, 5% B; t=0.35 min, 5% B; t=14.5 min, 66% B; t=18 min, 99.5% B; t=33.50 min, 99.5% B; t=35.51 min, 99.5% B. The stop time was 40 min. For MS, data were acquired by Agilent 6540 Q-TOF mass spectrometer (Agilent Technologies) operating in the positive and negative ion mode using Agilent Jet Stream ESI source. The capillary voltage was set at +3800 V (positive mode) and −3800 V (negative mode) with nozzle voltages of +0 V and −0 V, respectively. Other source conditions were kept constant in all the experiments as follows: gas temperature was kept constant at 300 ℃, drying gas (nitrogen) was set at the rate of 7 L/min, and the pressure of nebulizer gas (nitrogen) was 40 pounds per square inch. The sheath gas was kept at a flow rate of 10 L/min and was maintained at a temperature of 330 ℃. The voltages of the Fragmentor, Skimmer 1 and OctopoleRFPeak were 135 V, 65 V and 750 V, respectively. The scan range was adjusted to 80–1700 m/z at the acquisition rate of 2 spectra/s. MS/MS acquisition was operated in the same parameter as in MS acquisition. Collision Energy was used at 10, 20 or 40 eV for fragmentation of the targeted compounds.

Data processing and statistical data analysis

All mass spectral data were acquired using Agilent MassHunter Qualitative Analysis software (version B.05.00; Agilent Technologies). To optimize feature detection and discovery, two software packages: Mass Hunter Qualitative Analysis and open-source software XCMS (version 1.38.0) operating in R, which adopted different peak detection and alignment algorithms, were used.[30] For Mass Hunter Qualitative Analysis software,

data preprocessing including baseline correction, noise calculations and molecular features extraction were performed with built-in small molecule extraction algorithm. Data were subsequently processed using Mass Profiler Professional (Agilent Technologies) for peak alignment, data filtering and statistical analysis. For XCMS, raw data files were first converted to mzDATA format and peak detection were performed with centWave alogrithm in XCMS.[31] Data were subsequently processed using XCMS for peak alignment and data filtering. MetaboAnalyst 2.0 (http://www.metaboanalyst.ca)[32] was used for statistical analysis.[24] Further data processing including normalization, scaling and filtering were performed prior to statistical analysis in both software. Only variables that are present in at least 60% of any group and with intensity of at least 4.0×10^3 were included for analysis in order to reduce noise and low abundance metabolites. The MS data were $\log_2$-transformed and mean-centered with unit variance scaling for statistical analysis. Principle component analysis (PCA) and hierarchical clustering were performed for unsupervised multivariate statistical analysis. Partial-least squares discrimination analysis (PLS-DA) was performed as supervised method to identify important variables with discriminative power. PLS-DA models were validated based on multiple correlation coefficient (R^2) and cross-validated R^2 (Q^2) in cross-validation and permutation test by applying 2000 iterations ($P>0.001$). The significance of the biomarkers was ranked using the variable importance in projection (VIP) score (>1) from the PLS-DA model. For univariate analysis of candidate specific biomarkers in culture supernatant, statistical significance was determined using one-way ANOVA with Tukey's *post-hoc* test between multiple groups of MTB and individual NTM species (*M. avium* complex, *M. bovis* BCG, *M. chelonae*, *M. fortuitum* and *M. kansasii*), and Student's *t*-test for comparison between MTB and NTB. $P<0.05$ was considered to be statistically significant. Volcano plot with fold change >5 and $P<0.05$ was performed where appropriate. Box-whisker plots were produced using GraphPad Prism software (GraphPad Software Inc., San Diego, CA, USA). Extracted ion chromatograms of potential specific metabolites identified by statistical analysis were manually viewed to confirm the differences in peak areas between MTB and NTB samples. Metabolites were further filtered using CAMERA package in R, MassHunter and manual inspection to exclude possible fragments, dimers, adducts and isotopes.[33] Specific metabolites that were detected by both Mass Profiler Professional and MetaboAnalyst to be statistically significant were considered to be potential biomarkers.

Metabolite identification

MS/MS fragmentation was performed on the identified potential specific biomarkers. Identification of potential biomarkers was carried out by searching METLIN database (http://metlin.scripps.edu/), Human Metabolome Database (http://www.hmdb.ca/), *E. coli* Metabolome Database (http://www.ecmdb.ca/), MassBank (http://www.massbank.jp/), LipidMaps (http://www.lipidmaps.org/) and KEGG database (http://www.genome.jp/kegg) using exact molecular weights or MS/MS fragmentation pattern data and literature search. Mass accuracy tolerance within 10 ppm was used as the mass window for database search. For confirmation of metabolite identity using authentic chemical standard, MS/MS fragmentation pattern of chemical standard was compared with that of candidate metabolite under same LC–MS condition to reveal any matching. In case of unknown metabolites, molecular formulae were generated using Mass Profiler Professional (Agilent Technologies).

RESULTS

Visual inspection of total ion chromatograms

We characterized and compared the metabolomes of culture supernatants from nine MTB and nine NTM strains. The total ion chromatograms from the nine MTB strains shared considerable similarity, whereas significant differences were observed in the chromatograms obtained from different NTM species. Representative examples of chromatograms obtained from each species are shown in Supplementary Figure S1.

PCA and PLS-DA modeling

To compare the metabolomes between MTB and NTM strains, both multi- and univariate analyses were performed. For multivariate analysis, PCA showed that 50.1% of the total variance in the data was represented by the first two principal components (Figure 1A). The 2D-PCA score plot revealed that the MTB strains were closely related to each other and can be distinguished from the NTM strains based on the first two principal components, with the MTB strains clearly separated from NTM strains along principal component 1, which represented 40.3% of the variance. In view of the significant separation achieved using PCA, supervised analysis PLS-DA (Figure 1B) was subsequently performed to maximize the separation and identify additional metabolites to those identified using PCA. In the PLS-DA score plot, the separation between different *Mycobacterium* species is more prominent. Potential metabolites were selected based on the VIP score (>1). Hierarchical clustering analysis was performed based on the degree of similarity of metabolite abundance profiles to show the global overview of all culture supernatant metabolites detected (Figure 2). Metabolites with similar abundance patterns were positioned closer together. The heat map and dendrogram indicated the close clustering of the MTB strains and their separation from the NTM strains. To further confirm the specificity and significance of potential metabolites identified from PCA and PLS-DA, univariate analysis of each metabolite was performed using one-way ANOVA and Student's *t*-test. A total of 24 potential metabolites contributing most to the variation between MTB and NTM with significantly higher level in MTB strains were selected for further identification (Table 1).

Identification of potential biomarkers specific to *M. tuberculosis*

The 24 metabolites were identified by MS/MS fragmentation and their predicted molecular formulae are shown in Table 1. All metabolites except *m/z* 206.1388 were only found in MTB but not NTM strains. Seven (*m/z* 206.1388, *m/z* 483.2759, *m/z* 521.2498, *m/z* 524.3598, *m/z* 530.3698, *m/z* 540.3572 and *m/z* 582.3642) of the 24 metabolites were identified by MS/MS, while the other 17 metabolites represent potentially novel metabolites with no match against known compounds or databases (Table 1 and Figure 3).

The metabolite *m/z* 206.1388 was identified as dexpanthenol (metabolite NO 1 in Table 1) with molecular formula $C_9H_{19}NO_4$ by database searches in METLIN and Massbank, and confirmed by MS/MS using commercially available authentic chemical standard of dexpanthenol (Figure 4). Although it was found in MTB strains with significantly higher level, low levels of *m/z* 206.1388 (at approximately 10- to 100-fold lower levels) were also detected in *M. avium* complex, *M. bovis* BCG, *M. chelonae*, *M. fortuitum* and *M. kansasii* (Figure 3).

One metabolite *m/z* 521.2498 was identified as a tetrapeptide, Val-His-Glu-His (metabolite NO 17 in Table 1) with molecular formula $C_{22}H_{32}N_8O_7$ in METLIN. Some bacteria may produce short peptides as pheromones, which are involved in quorum-sensing. However, the significance of this short peptide from MTB culture supernatants remains to be determined.

Four metabolites, *m/z* 524.3598, *m/z* 530.3698, *m/z* 540.3572 and *m/z* 582.3642 (metabolite NOs 18, 19, 20 and 21 in Table 1), were identified as 1-tuberculosinyladenosine (1-TbAd) or its derivatives by MS/MS. The metabolite *m/z* 540.3572 was identified as 1-TbAd with molecular formula $C_{30}H_{46}N_5O_4$. MS/MS fragmentation showed

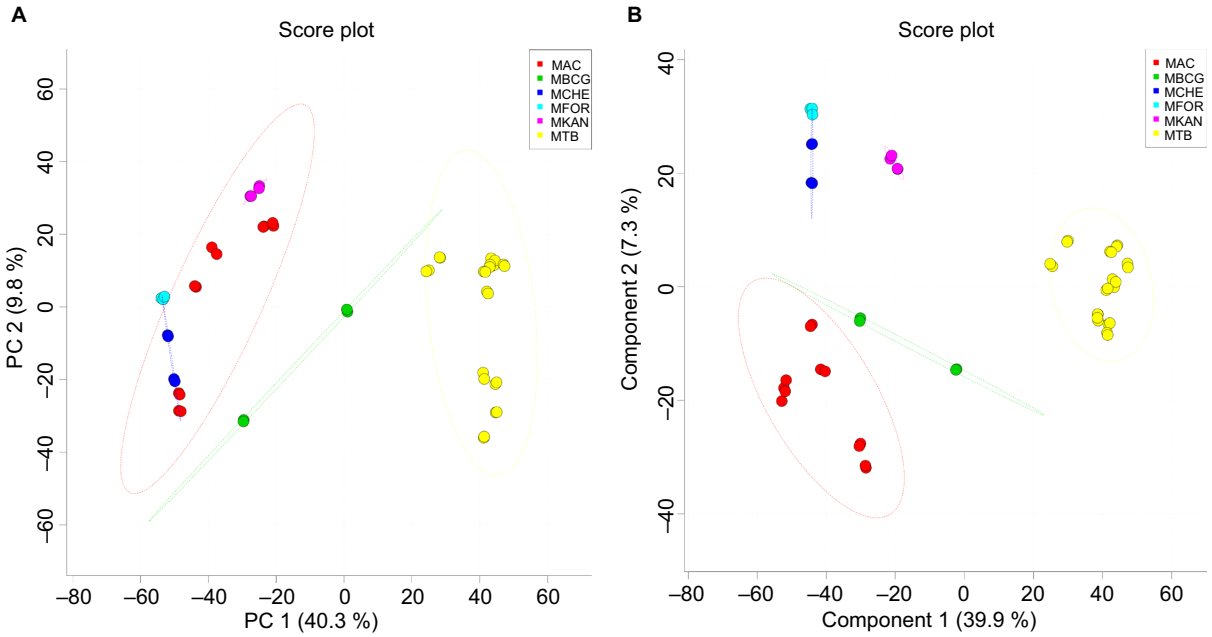

Figure 1 (**A**) PCA score plot and (**B**) PLS-DA score plot generated using MetaboAnalyst based on culture supernatant in positive mode. PLS-DA models were validated using R^2 and Q^2 based on LOOCV. Five-component model was selected as optimized model with R^2=0.99 and Q^2=0.99. The significance of the model was demonstrated by permutation test with 2000 testing iterations using separation distance and P value <0.001 was obtained. LOOCV, leave one out cross-validation; MAC, *M. avium* complex; MBCG, *M. bovis* BCG; MCHE, *M. chelonae*; MFOR, *M. fortuitum*; MKAN, *M. kansasii*.

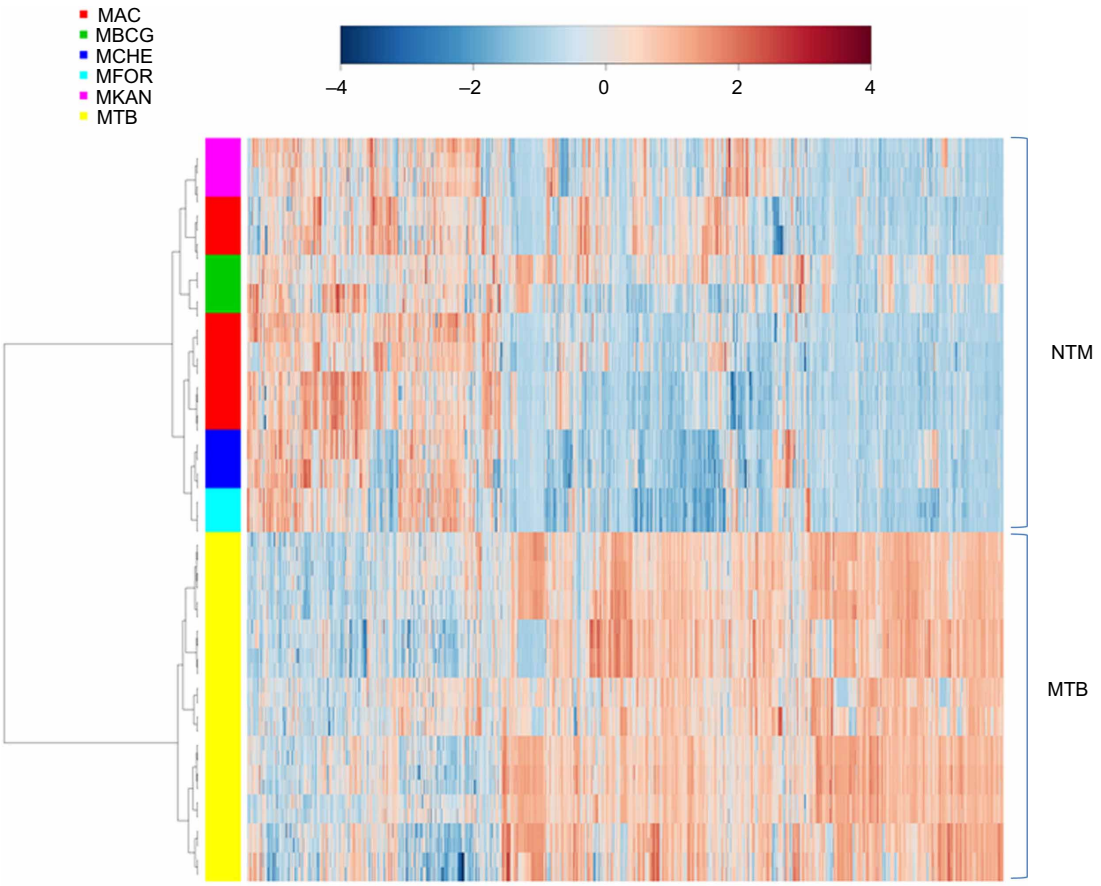

Figure 2 Hierarchical clustering analysis generated using MetaboAnalyst based on culture supernatant in positive mode. Each bar represented a metabolite colored by its abundance intensities on normalized scale from blue (decreased level) to red (increased level). The dendrogram on the left was constructed based on the metabolite abundance profiles. MAC, *M. avium* complex; MBCG, *M. bovis* BCG; MCHE, *M. chelonae*; MFOR, *M. fortuitum*; MKAN, *M. kansasii*.

Table 1 Specific metabolites in culture supernatant of *M. tuberculosis*

Metabolite NO	*m/z*	Retention time (min)	Ionization mode	Ion	MS/MS fragment masses	*P* value[a]	VIP score[b]	Molecular formula	Putative identity
1	206.1388	2.30	Positive	$[M+H]^+$	58.0667, 76.0768, 118.0666, 170.1205, 188.1311	<0.001	1.26	$C_9H_{19}NO_4$	Dexpanthenol[c]
2	364.1234	11.33	Positive	$[M+H]^+$	187.0388, 200.0507, 215.0329, 230.0578, 245.0810	<0.001	1.10	$C_{18}H_{21}NO_{5S}$	No match
3	251.1277	13.10	Positive	$[M+H]^+$	180.0771, 191.1055, 194.0906, 208.1088, 251.1256	<0.001	1.38	$C_{14}H_{18}O_4$	No match
4	306.135	6.94	Negative	$[M-H]^-$	170.061, 187.0621, 191.0583, 200.1065, 215.0957, 259.0858, 274.1086	<0.001	1.85	$C_{16}H_{21}NO_5$	No match
5	321.2432	16.49	Negative	$[M-H]^-$	59.0141, 321.2435	<0.001	1.73	$C_{20}H_{34}O_3$	No match
6	336.1441	8.86[d]	Positive	$[M+H]^+$	101.0610, 129.0558, 258.1125, 272.0927, 304.1194, 336.1459	<0.001	1.38	$C_{17}H_{21}NO_6$	No match
7	322.1285	6.85	Positive	$[M+H]^+$	148.0754, 176.0690, 208.0978, 244.0972, 272.0938, 290.1040, 322.1051	<0.001	1.24	$C_{16}H_{19}NO_6$	No match
8	324.1809	7.68	Positive	$[M+H]^+$	150.0910, 151.0743, 193.0859	<0.001	1.32	$C_{17}H_{25}NO_5$	No match
9	350.16[d]	9.49	Positive	$[M+H]^+$	111.0446, 143.0704, 286.1081, 318.1342, 350.1605	<0.001	1.30	$C_{18}H_{23}NO_6$	No match
10	352.1407	5.00	Negative	$[M-H]^-$	201.0765, 216.1060, 260.0936, 352.1399	<0.001	1.84	$C_{17}H_{23}NO_7$	No match
11	364.1758	10.30	Positive	$[M+H]^+$	69.0681, 97.0628, 111.0784, 125.0572, 148.073, 176.0682, 228.1363, 254.1158, 256.1297	<0.001	1.33	$C_{19}H_{25}NO_6$	No match
12	368.1698	5.39	Positive	$[M+H]^+$	148.0810, 176.0707, 268.1120, 276.1222, 350.1586, 368.1871	<0.001		$C_{18}H_{25}NO_7$	No match
13	372.2186	14.37	Negative	$[M-H]^-$	131.0375, 298.2178, 342.1916	<0.001	1.30	$C_{22}H_{31}NO_4$	No match
14	380.1720[d]	5.92	Negative	$[M-H]^-$	59.014, 91.0408, 170.0606, 229.108, 244.1331, 273.101, 288.1241	<0.001	1.62	$C_{19}H_{27}NO_7$	No match
15	404.2793	17.74	Positive	$[M+H]^+$	109.1002, 148.0852, 176.0703, 208.0991, 372.2550, 386.2692, 404.2837	<0.001	1.2867	$C_{24}H_{37}NO_4$	No match
16	410.1812	7.42	Positive	$[M+H]^+$	148.0760, 286.1077, 318.1336, 392.1708	<0.001	1.24	$C_{20}H_{27}NO_8$	No match
17	521.2498	12.23	Positive	$[M+H]^+$	176.0700, 209.0964, 236.0915, 252.1225, 375.1895, 482.2194, 472.1969, 489.2208, 521.2481	<0.001	1.29	$C_{22}H_{32}N_8O_7$	Val-his-glu-his
18	524.3598	14.62	Positive	$[M]^+$	136.0624, 163.1480, 252.1101, 273.2564, 408.3139, 524.3611	<0.001	1.41	$C_{30}H_{46}N_5O_3$	1-tuberculosinyl2′-deoxyadenosine
19	530.3698	14.95	Positive	$[M]^+$	109.0512, 126.0776, 149.1324, 163.1480, 177.1635, 241.0940, 258.1198, 273.2580, 398.3279, 530.3711	<0.001	1.41	$C_{29}H_{48}N_5O_4$	1-tuberculosinyl derivative
20	540.3572[d]	14.18	Positive	$[M]^+$	136.0633, 163.1499, 268.1063, 273.2599 408.3145, 540.3557	<0.001	1.32	$C_{30}H_{46}N_5O_4$	1-tuberculosinyladenosine
21	582.3642	15.27	Positive	$[M]^+$	136.0621, 163.1481, 273.2580, 310.1137, 408.3105, 582.3660	<0.001	1.41	$C_{32}H_{48}N_5O_5$	1-tuberculosinyl-O-acetyladenosine
22	617.2444	7.51	Negative	$[M-H]^-$	59.0144, 151.0400, 311.1136, 421.1501, 439.1595, 617.2444	<0.001	1.75	$C_{30}H_{34}N_8O_7$	No match
23	388.2473	16.78	Positive	$[M+H]^+$	148.0755, 163.1482, 176.0706, 208.0965, 356.2231, 388.2493	<0.001	1.27	$C_{23}H_{33}NO_4$	No match
24	483.2759	14.05	Negative	$[M-H]^-$	152.9966, 227.0336, 255.2329	<0.001	1.7 2	$C_{22}H_{45}O_9P$	PG(16 : 0/0 : 0)

Abbreviation: PG, phosphatidylglycerol.
[a] *P* value from ANOVA analysis.
[b] VIP score based on PLS-DA. VIP score >1 is considered to be statistically significant.
[c] Confirmed by MS/MS fragmentation pattern matching with commercially available authentic chemical standard.
[d] Detected in both positive and negative mode.

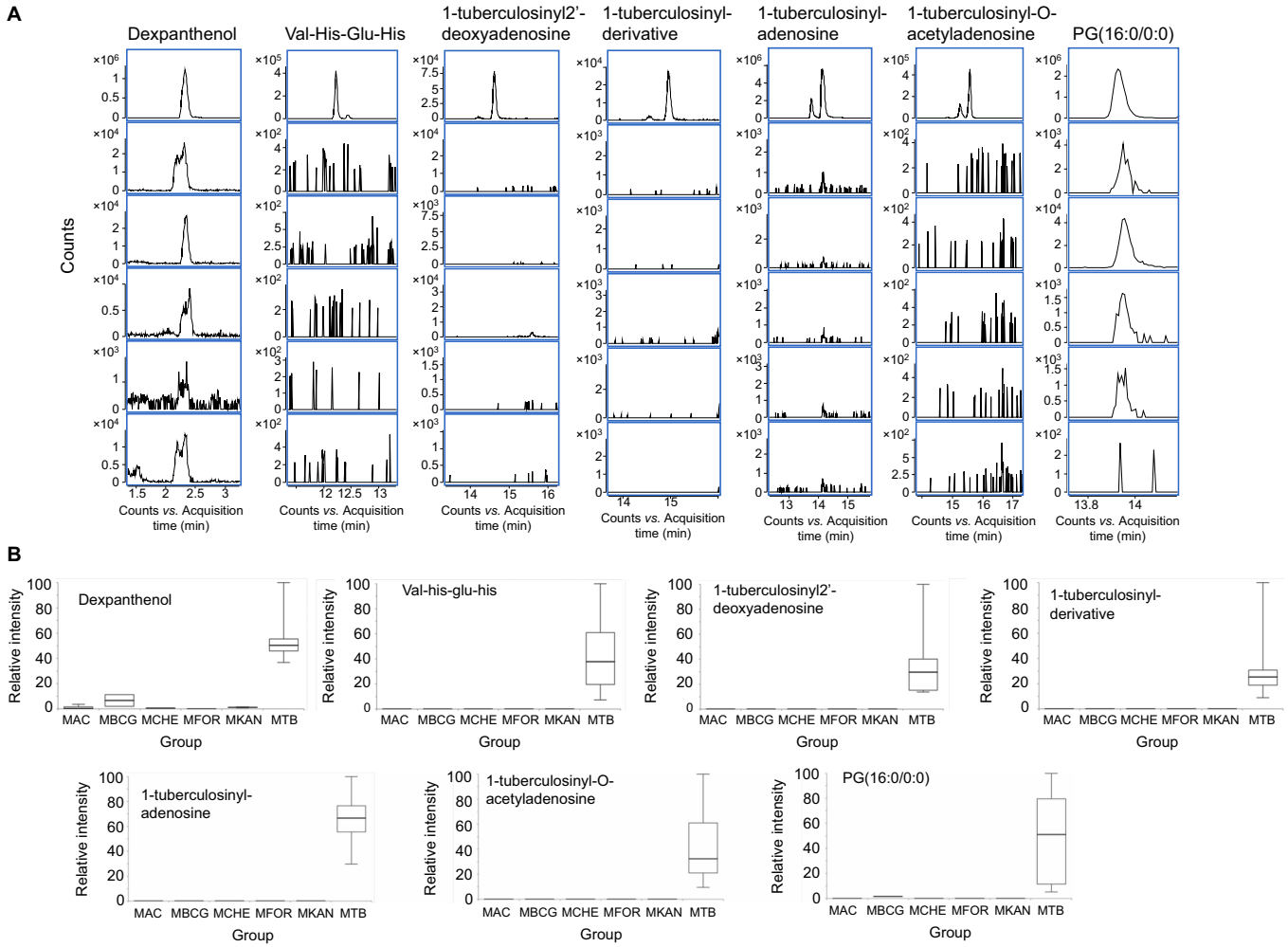

Figure 3 (**A**) Extracted ion chromatogram and (**B**) box-whisker plots of dexpanthenol, Val-His-Glu-His, 1-tuberculosinyl 2′-deoxyadenosine, 1-tuberculosinyl-derivative, 1-tuberculosinyl-adenosine, 1-tuberculosinyl-O-acetyladenosine and PG (16 : 0/0 : 0). MAC, *M. avium* complex; MBCG, *M. bovis* BCG; MCHE, *M. chelonae*; MFOR, *M. fortuitum*; MKAN, *M. kansasii*.

major peaks at *m/z* 136.0616, *m/z* 268.1041, *m/z* 273.2575 and *m/z* 408.3127, compatible with MS/MS profiles of 1-TbAb (Figure 5A).[34] The metabolite *m/z* 524.3598 also possessed major peaks at *m/z* 136.0615, *m/z* 273.2571 and *m/z* 408.3128. However, instead of *m/z* 268.1041 ($C_{10}H_{14}N_5O_4$), a peak at *m/z* 252.1088 ($C_{10}H_{14}N_5O_3$) was detected, suggesting that *m/z* 524.3598 is likely a metabolite of 1-TbAd, 1-tuberculosinyl 2′-deoxyadenosine, which is predicted to have a hydrogen atom in place of the hydroxyl group in C2′ of the ribose to give 2′-deoxyribose (Figure 5B). The metabolite *m/z* 582.3642 also possessed major peaks at *m/z* 136.0609, *m/z* 273.2569 and m/z 408.3115. However, instead of *m/z* 268.1041 ($C_{10}H_{14}N_5O_4$), a peak at *m/z* 310.1141 ($C_{12}H_{16}N_5O_5$) was detected, suggesting that *m/z* 582.3642 is also a metabolite of 1-TbAd, 1-tuberculosinyl-O-acetyladenosine, which is predicted to have an acetyl-group replacing the hydrogen atom of a hydroxyl group in O-2 position of the ribose (Figure 5C). Another tuberculosinyl-derivative was found to have molecular cation at *m/z* 530.3698 with molecular formula $C_{29}H_{48}N_5O_4$. Its corresponding MS/MS spectrum contains fragments at *m/z* 273.2610 referring to the polyunsaturated C20 hydrocarbon ($C_{20}H_{33}^+$), *m/z* 398.2908 representing to the peak ($C_{24}H_{40}N_5^+$) after loss of ribose from the derivative, *m/z* 258.1219 corresponding to the peak ($C_9H_{16}N_5O_4^+$) after loss of the polyunsaturated C20 hydrocarbon and *m/z* 126.0787 ($C_4H_8N_5^+$) referring to the base replacing the adenine (Figure 5D).

Another metabolite *m/z* 483.2759 was identified as a monoacylglycerophosphoglycerol, PG (16 : 0/0 : 0) or 1-hexadecanoyl-sn-glycero-3-phospho-(1′-sn-glycerol) (metabolite NO 24 in Table 1), with molecular formula $C_{22}H_{45}O_9P$ in METLIN (Figure 5E). PG (16 : 0/0 : 0) has been reported in the lipidomic database, Mtb LipidDB, of MTB, being found in the mycobacterial cell extracts.[35] PG, a subclass of glycerophospholipid, is found in bacterial membranes where PG is utilized as a precursor for cardiolipin synthesis.[36] It is possible that this metabolite, PG (16 : 0/0 : 0), is a unique component in the cell membrane of MTB not found in other mycobacteria. However, it remains to be determined whether PG (16 : 0/0 : 0) may be involved in virulence or other functions in MTB.

DISCUSSION

Using metabolomics approach, we identified specific metabolites in culture supernatant of MTB. As these extracellular metabolites are either secreted or released from cell wall components of MTB, they may be present in the circulating blood or other body fluids of infected patients, and hence may represent potential biomarkers for non-invasive diagnosis of tuberculosis. The exclusion of metabolites

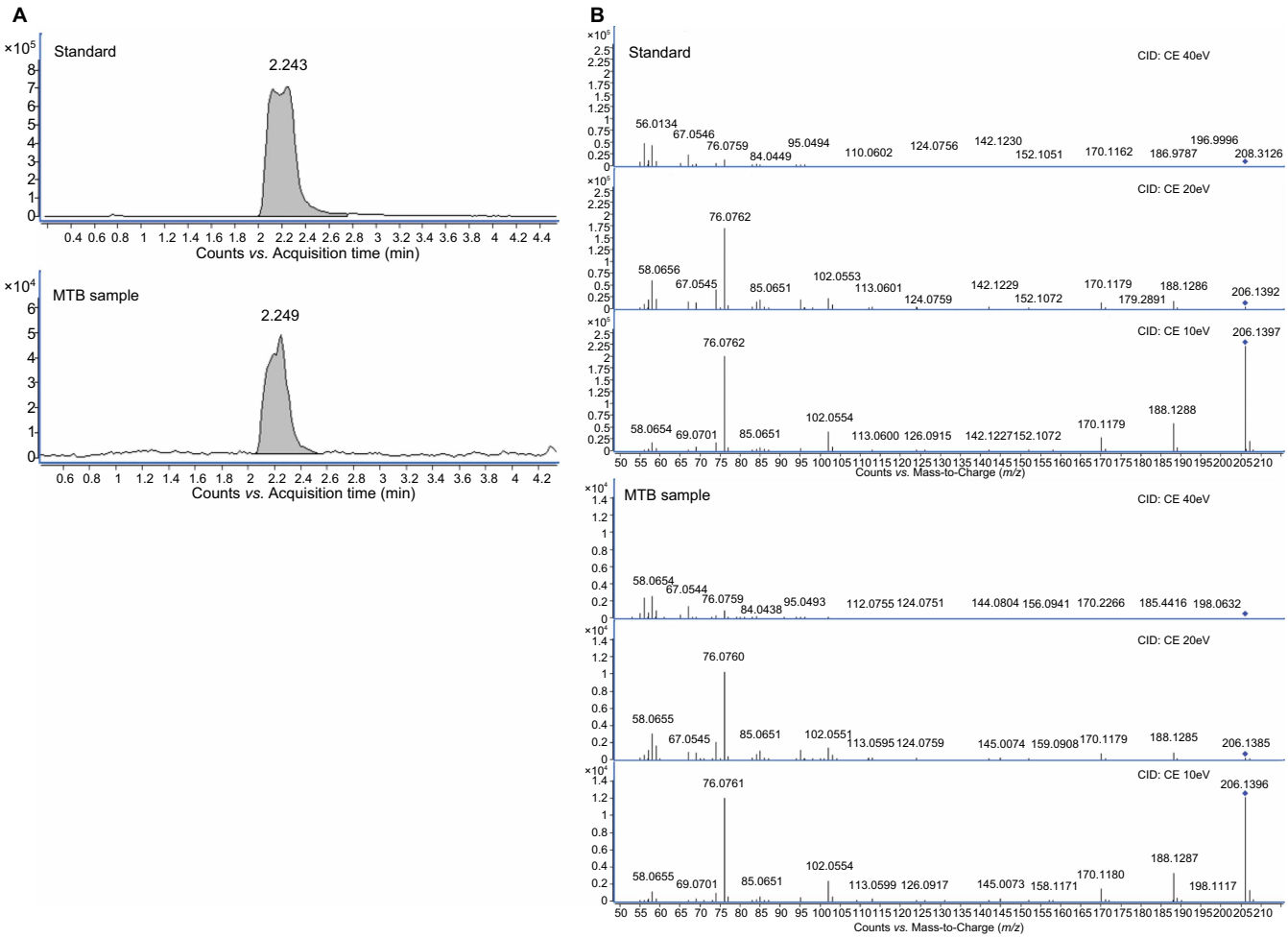

Figure 4 (**A**) Extracted ion chromatogram and (**B**) MS/MS mass spectra of dexpanthenol standard and *m/z* 206.1388 in MTB sample. MS/MS fragmentations have been performed at 10, 20 and 40 eV. CID, collision-induced dissociation; CE, collision energy.

present in NTM species was important, since these NTM species can also cause disseminated infections mimicking tuberculosis in immunocompromised hosts, which are also difficult to diagnose.[7] In this study, 24 metabolites with significantly higher levels in MTB were identified. Many of the metabolites were unidentified by MS/MS against database matching, suggesting that these are potential novel compounds. This is not unexpected, since metabolomics for the study of microbes including mycobacteria is still an emerging field and the number of known metabolites from MTB is very limited. Further studies on the chemical structure and biosynthetic pathway of these potential novel metabolites would help understand their biological function in MTB. As the present results were based on *in vitro* data obtained from cultures using Middlebrook 7H9 medium only, the significance of these metabolites in different growth phases and *in vitro* and *in vivo* environments should also be explored. More importantly, metabolomics studies on clinical samples from patients with tuberculosis are warranted to explore for the presence of these MTB-specific extracellular metabolites and their potential role as diagnostic biomarkers.

The presence of high levels of dexpanthenol in the culture supernatant of MTB strains is intriguing. Dexpanthenol is the alcohol analog of pantothenic acid (also called vitamin B5) which is used for the synthesis of Coenzyme A, an important cofactor in central metabolism. While the biosynthetic pathway of pantothenic acid is present in MTB,[37] production of dexpanthenol in bacteria has not been reported previously. Nevertheless, it has been demonstrated that dexpanthenol can act as a substrate for pantothenate kinase in MTB to produce 4′-phosphopantothenol which can inhibit the activity of 4′-phosphopantothenoylcysteine synthase and 4′-phosphopantothenoylcysteine decarboxylase and eventually affect the biosynthesis of Coenzyme A.[38] Moreover, dexpanthenol has been reported to exhibit antimicrobial activity against some bacteria and protozoa, especially those that are auxotrophic for pantothenate.[39,40] The much higher levels of dexpanthenol produced by MTB strains than NTM strains may suggest a possible role in survival benefit through inhibition of other bacteria. Further studies using ^{13}C carbon source in the media may help confirm dexapanthenol and other metabolites are true metabolic product of MTB and to examine the function and biosynthesis pathway.

Apart from 1-TbAd, we also identified three previously undescribed derivatives of 1-TbAd in the culture supernatant of MTB. 1-TbAd is a recently identified amphiphilic diterpene-linked adenosine found in the lipid extracts of culture supernatant of MTB but absent in those of *M. bovis* BCG.[34] The production of 1-TbAd requires the virulence-associated enzyme Rv3378c.[34] Since *Rv3377c–Rv3378c* locus has also been shown to be essential for optimal phagosome maturation arrest, it may suggest a role of 1-TbAd in phagosome survival.[41] Our data confirmed the specificity of 1-TbAd which can be found only in the culture supernatant of MTB but not in those of NTM strains, thus

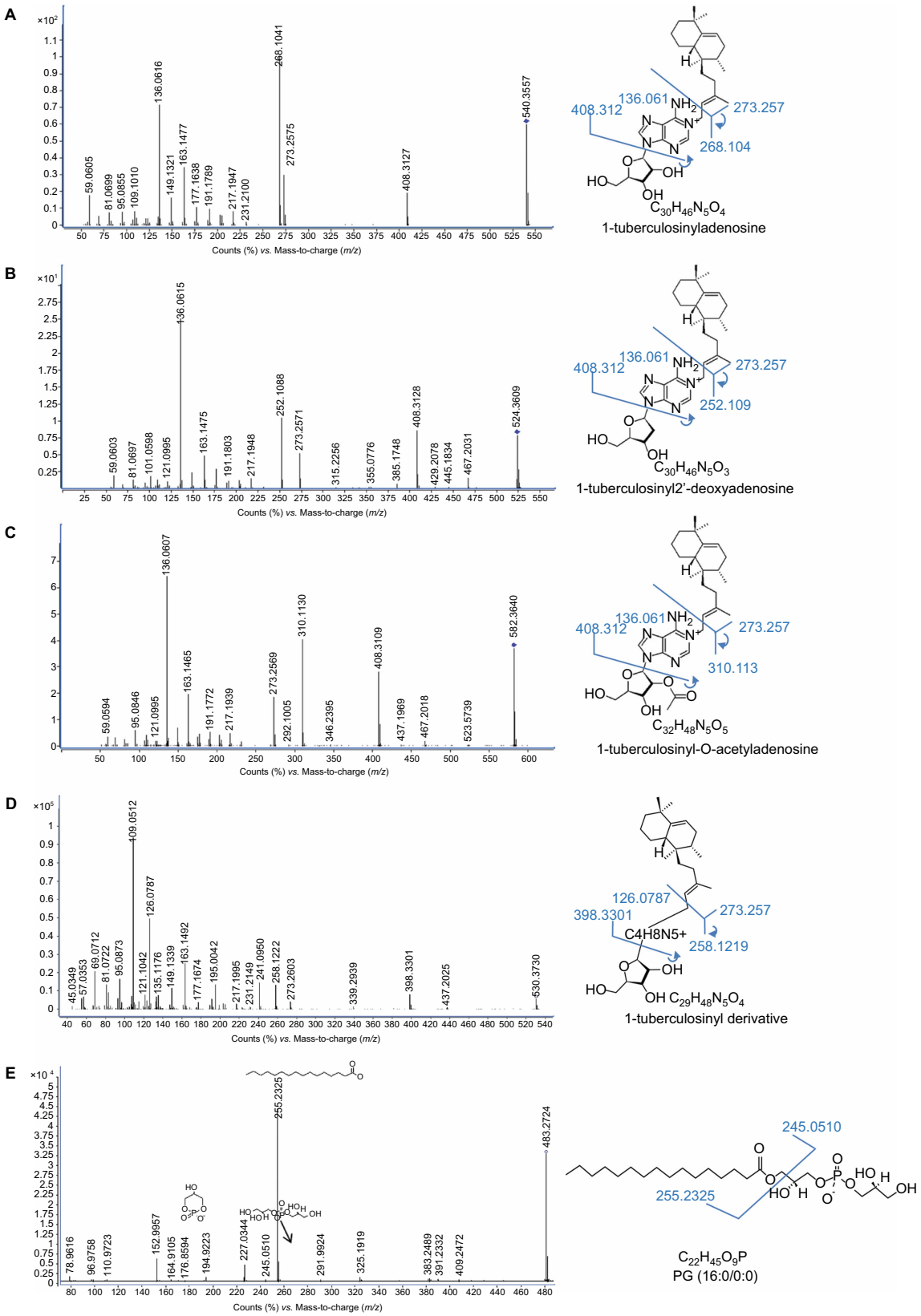

Figure 5 MS/MS mass spectra and predicted structures of (**A**) *m/z* 540.3572, (**B**) *m/z* 524.3598, (**C**) *m/z* 582.3642, (**D**) *m/z* 530.3698 and (**E**) *m/z* 483.2759 in MTB samples. MS/MS fragmentations performed at 20 eV were shown. The hydrogen group replacing the hydroxyl-group in the ribose of *m/z* 524.3598, and the acetyl-group replacing the hydroxyl-group in the ribose of *m/z* 582.3642 may be present in C2 or C3 position. The base of *m/z* 530.3698 is undetermined.

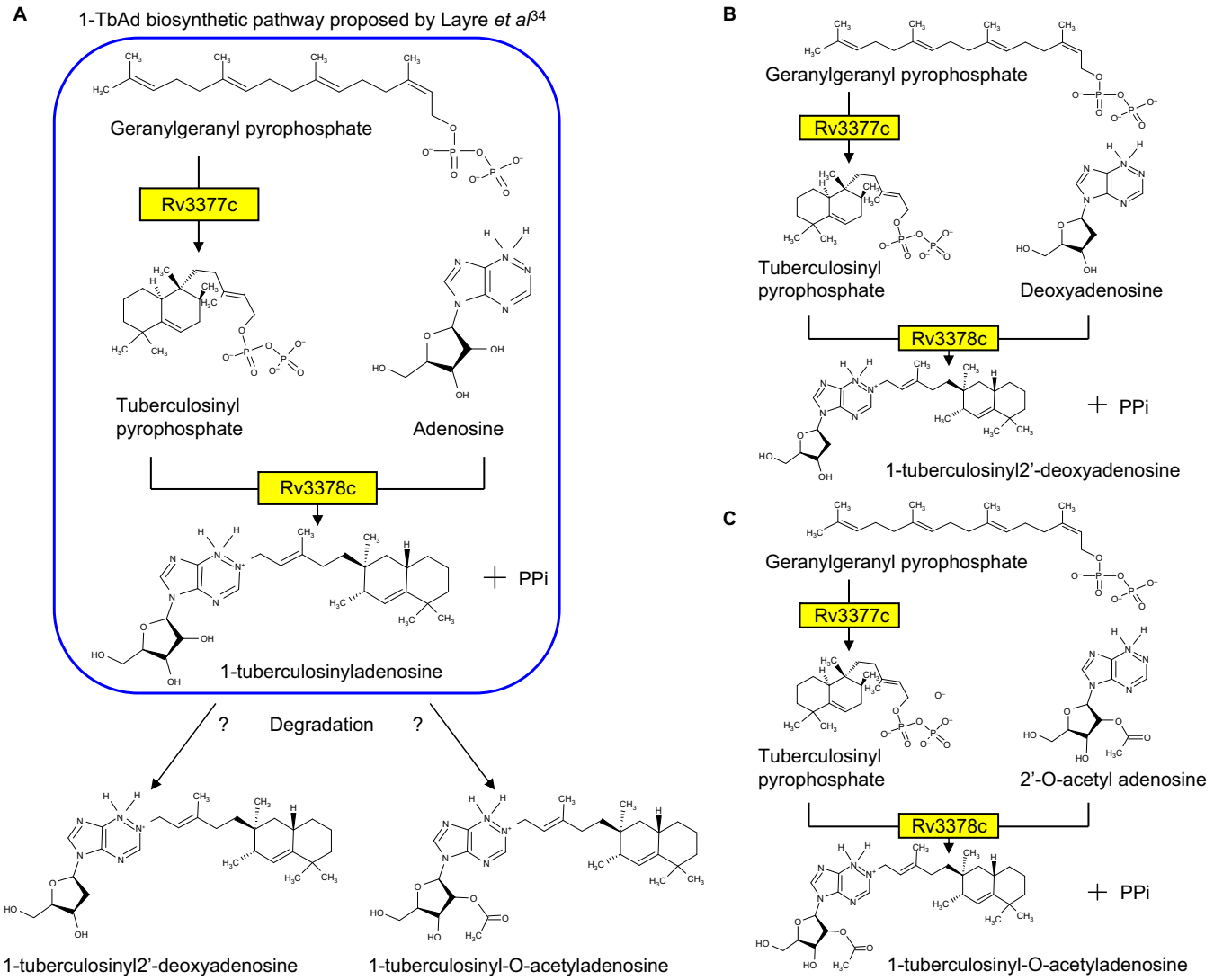

Figure 6 Proposed biosynthetic pathways of the two derivatives of 1-tuberculosinyl-adenosine: 1-tuberculosinyl 2′-deoxyadenosine and 1-tuberculosinyl-O-acetyladenosine based on the 1-TbAd biosynthetic pathway proposed by Layre *et al.*[34] (**A**) Hypothesis 1: the two derivatives are the degradation products of 1-TbAd. Hypothesis 2: the virulence associated enzyme, Rv3378c, catalyzes the conversion of (**B**) tuberculosinyl pyrophosphate and deoxyadenosine to 1-tuberculosinyl 2′-deoxyadenosine and (**C**) tuberculosinyl pyrophosphate and 2′-O-acetyl adenosine to 1-tuberculosinyl-O-acetyladenosine.

supporting 1-TbAd as a potential virulence factor in MTB. Three derivatives of 1-TbAd, 1-tuberculosinyldeoxyadenosine, 1-tuberculosinyl-O-acetyladenosine and another novel derivative with undetermined base, with approximately 10- to 100-fold lower levels than 1-TbAd, were also identified in the culture supernatant of MTB. 1-tuberculosinyldeoxyadenosine and 1-tuberculosinyl-O-acetyladenosine likely differ from 1-TbAd in the ribose group of the nucleoside. The remaining 1-tuberculosinyl derivative likely differs from 1-TbAd in the base, but the base structure remains to be ascertained. The two derivatives, 1-tuberculosinyldeoxyadenosine and 1-tuberculosinyl-O-acetyladenosine, are either degradation products of 1-TdAd (Figure 6A) or synthesized by Rv3378c similar to the previously proposed biosynthetic pathway for of 1-TdAd (Figures 6B and 6C).[34] Rv3378c is believed to have combined phosphatase and tuberculosinyl transferase activity, catalyzing the conversion of tuberculosinyl pyrophosphate and adenosine to 1-TbAd, using adenosine as the nucleophilic substrate.[34] Therefore, it may also potentially catalyze the conversion of tuberculosinyl pyrophosphate and deoxyadenosine to 1-tuberculosinyl 2′-deoxyadenosine (Figure 6B), and similarly, the conversion of tuberculosinyl pyrophosphate and 2′-O-acetyl adenosine to 1-tuberculosinyl-O-acetyladenosine (Figure 6C). While it is unclear if 2′-O-acetyl adenosine exists as a metabolite in MTB, deoxyadenosine, a derivative of adenosoine, presents in all living organisms. As the present results were obtained from MS/MS analysis only, further studies using nuclear magnetic resonance may help elucidate the exact structures and biosynthetic pathways of these metabolites and the molecular mechanisms of virulence of 1-TbAd and/or its deriviatives in MTB.

Metabolomics is an uprising tool in microbiology and infectious disease research, providing a revolutionary method to study both the pathogen itself and the host response to the infection. For tuberculosis, the technique has been used to identify various *Mycobacterium* species and carbon utilization of MTB, and characterize metabolites of hyper-virulent strains.[13,14,23] These findings may help better understand the biology and virulence factors of MTB. On the other hand, metabolomics applied on direct patient samples may reveal metabolites generated during infection of the host, which can provide insights on the diagnosis, pathogenesis

and host–pathogen interactions. For mycobacteria, it has been shown that volatile organic compounds in the urine samples from tuberculosis patients can be distinguished from those from healthy subjects.[24] Moreover, a study using nuclear magnetic resonance spectroscopy-based metabolomics showed that sera of tuberculosis patients can be distinguished from those of healthy controls.[25] Therefore, metabolomics approach may help identify potential biomarkers for diagnosis of tuberculosis. In another study using serum metabolomics approach on leprosy patients, higher levels of polyunsaturated fatty acids were found among patients having higher bacterial indices, which may provide clues on the biological pathways involved in the immunomodulation of leprosy.[42] With the increasing applications of metabolomics technology on both microbial and clinical samples from patients with appropriate controls, we expect to witness a gross expansion of our knowledge on microbial metabolites, including the discovery of novel metabolites and potential biomarkers for diagnosis of infections such as tuberculosis.

ACKNOWLEDGEMENTS

This work is partly supported by the Hong Kong Special Administrative Region Research Fund for the Control of Infectious Diseases (Commissioned Study HK-09-01-10) of the Health, Welfare and Food Bureau; University Development Fund and Strategic Research Fund, The University of Hong Kong; and donation from Ms Eunice Lam on emerging infectious disease research.

1 World Health Organization. *Global tuberculosis report* 2014. Geneva: WHO, 2014. Available at http://www.who.int/tb/publications/global_report/en/ (accessed 2 December 2014).
2 Lai CC, Liu WL, Tan CK *et al*. Differences in drug resistance profiles of *Mycobacterium tuberculosis* isolates causing pulmonary and extrapulmonary tuberculosis in a medical centre in Taiwan, 2000–2010. *Int J Antimicrob Agents* 2011; **38**: 125–129.
3 Jiang Y, Dou X, Zhang W *et al*. Genetic diversity of antigens Rv2945c and Rv0309 in *Mycobacterium tuberculosis* strains may reflect ongoing immune evasion. *FEMS Microbiol Lett* 2013; **347**: 77–82.
4 Lin H, Ding H, Guo FB, Huang J. Prediction of subcellular location of mycobacterial protein using feature selection techniques. *Mol Divers* 2010; **14**: 667–671.
5 Chan JF, Trendell-Smith NJ, Chan JC *et al*. Reactive and infective dermatoses associated with adult-onset immunodeficiency due to anti-interferon-gamma autoantibody: Sweet's syndrome and beyond. *Dermatology* 2013; **226**: 157–166.
6 Doucette K, Fishman JA. Nontuberculous mycobacterial infection in hematopoietic stem cell and solid organ transplant recipients. *Clin Infect Dis* 2004; **38**: 1428–1439.
7 Lau SK, Curreem SO, Ngan AH, Yeung CK, Yuen KY, Woo PC. First report of disseminated *Mycobacterium* skin infections in two liver transplant recipients and rapid diagnosis by hsp65 gene sequencing. *J Clin Microbiol* 2011; **49**: 3733–3738.
8 Getahun H, Harrington M, O'Brien R, Nunn P. Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resourceconstrained settings: informing urgent policy changes. *Lancet* 2007; **369**: 2042–2049.
9 Marais BJ, Pai M. New approaches and emerging technologies in the diagnosis of childhood tuberculosis. *Paediatr Respir Rev* 2007; **8**: 124–133.
10 Cheng VC, Yam WC, Hung IF *et al*. Clinical evaluation of the polymerase chain reaction for the rapid diagnosis of tuberculosis. *J Clin Pathol* 2004; **57**: 281–285.
11 Moore DF, Curry JI. Detection and identification of *Mycobacterium tuberculosis* directly from sputum sediments by Amplicor PCR. *J Clin Microbiol* 1995; **33**: 2686–2691.
12 Olivier I, Loots DT. An overview of tuberculosis treatments and diagnostics. What role could metabolomics play? *J Cell Tissue Res* 2011; **11**: 2655–2671.
13 Olivier I, Loots DT. A metabolomics approach to characterise and identify various *Mycobacterium* species. *J Microbiol Methods* 2012; **88**: 419–426.
14 Meissner-Roloff RJ, Koekemoer G, Warren RM, Loots DT. A metabolomics investigation of a hyper- and hypo-virulent phenotype of Beijing lineage *M. tuberculosis*. *Metabolomics* 2012; **8**: 1194–1203.
15 Schoeman JC, Loots DT. Improved disease characterisation and diagnostics using metabolomics: a review. *J Cell Tissue Res* 2011; **11**: 2673–2683.
16 Tam EW, Chen JH, Lau EC *et al*. Misidentification of *Aspergillus nomius* and *Aspergillus tamarii* as *Aspergillus flavus*: characterization by internal transcribed spacer, β-tubulin, and calmodulin gene sequencing, metabolic fingerprinting, and matrix-assisted laser desorption ionization–time of flight mass spectrometry. *J Clin Microbiol* 2014; **52**: 1153–1160.
17 To KK, Fung AM, Teng JL *et al*. Characterization of a *Tsukamurella* pseudo-outbreak by phenotypic tests, 16S rRNA sequencing, pulsed-field gel electrophoresis, and metabolic footprinting. *J Clin Microbiol* 2013; **51**: 334–338.
18 Woo PC, Lam CW, Tam EW *et al*. First discovery of two polyketide synthase genes for mitorubrinic acid and mitorubrinol yellow pigment biosynthesis and implications in virulence of *Penicillium marneffei*. *PLoS Negl Trop Dis* 2012; **6**: e1871.
19 Mahadevan S, Shah SL, Marrie TJ, Slupsky CM. Analysis of metabolomic data using support vector machines. *Anal Chem* 2008; **80**: 7562–7570.
20 Van QN, Klose JR, Lucas DA *et al*. The use of urine proteomic and metabonomic patterns for the diagnosis of interstitial cystitis and bacterial cystitis. *Dis Markers* 2003– 2004; **19**: 169–183.
21 Lv H, Hung CS, Chaturvedi KS, Hooton TM, Henderson JP. Development of an integrated metabolomic profiling approach for infectious diseases research. *Analyst* 2011; **136**: 4752–4763.
22 Fend R, Kolk AH, Bessant C, Buijtels P, Klatser PR, Woodman AC. Prospects for clinical application of electronic-nose technology to early detection of *Mycobacterium tuberculosis* in culture and sputum. *J Clin Microbiol* 2006; **44**: 2039–2045.
23 de Carvalho LP, Fischer SM, Marrero J, Nathan C, Ehrt S, Rhee KY. Metabolomics of *Mycobacterium tuberculosis* reveals compartmentalized co-catabolism of carbon substrates. *Chem Biol* 2010; **17**: 1122–1131.
24 Banday KM, Pasikanti KK, Chan EC *et al*. Use of urine volatile organic compounds to discriminate tuberculosis patients from healthy subjects. *Anal Chem* 2011; **83**: 5526–5534.
25 Zhou A, Ni J, Xu Z *et al*. Application of ^{1}H NMR spectroscopy-based metabolomics to sera of tuberculosis patients. *J Proteome Res* 2013; **12**: 4642–4649.
26 Somashekar BS, Amin AG, Tripathi P *et al*. Metabolomic signatures in guinea pigs infected with epidemic-associated W-Beijing strains of *Mycobacterium tuberculosis*. *J Proteome Res* 2012; **11**: 4873–4884.
27 Cha D, Cheng D, Liu M, Zeng Z, Hu X, Guan W. Analysis of fatty acids in sputum from patients with pulmonary tuberculosis using gas chromatography–mass spectrometry preceded by solid-phase microextraction and post-derivatization on the fiber. *J Chromatogr A* 2009; **1216**: 1450–1457.
28 Versalovic J, Carroll KC, Funke G, Jorgensen JH, Landry ML, Warnock DW. *Manual of clinical microbiology*. 10th ed. Washington, DC: ASM Press, 2011.
29 Tufariello JM, Jacobs WR Jr, Chan J. Individual *Mycobacterium tuberculosis* resuscitation-promoting factor homologues are dispensable for growth *in vitro* and *in vivo*. *Infect Immun* 2004; **72**: 515–526.
30 Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. *Anal Chem* 2006; **78**: 779–787.
31 Tautenhahn R, Böttcher C, Neumann S. Highly sensitive feature detection for high resolution LC/MS. *BMC Bioinformatics* 2008; **9**: 504.
32 Xia J, Mandal R, Sinelnikov IV, Broadhurst D, Wishart DS. MetaboAnalyst 2.0—a comprehensive server for metabolomics data analyst. *Nucleic Acids Res* 2012; **40**: W127–W133.
33 Kuhl C, Tautenhahn R, Böttcher C, Larson TR, Neumann S. CAMERA: an integrated strategy for compound spectra extraction and annotation of liquid chromatography/mass spectrometry data sets. *Anal Chem* 2012; **84**: 283–289.
34 Layre E, Lee HJ, Young DC *et al*. Molecular profiling of *Mycobacterium tuberculosis* identifies tuberculosinyl nucleoside products of the virulence-associated enzyme Rv3378c. *Proc Natl Acad Sci USA* 2014; **111**: 2978–2983.
35 Sartain MJ, Dick DL, Rithner CD, Crick DC, Belisle JT. Lipidomic analyses of *Mycobacterium tuberculosis* based on accurate mass measurements and the novel ''Mtb LipidDB''. *J Lipid Res* 2011; **52**: 861–872.
36 Schlame M. Cardiolipin synthesis for the assembly of bacterial and mitochondrial membranes. *J Lipid Res* 2008; **49**: 1607–1620.
37 Sambandamurthy VK, Wang X, Chen B *et al*. A pantothenate auxotroph of *Mycobacterium tuberculosis* is highly attenuated and protects mice against tuberculosis. *Nat Med* 2002; **8**: 1171–1174.
38 Kumar P, Chhibber M, Surolia A. How pantothenol intervenes in Coenzyme-A biosynthesis of *Mycobacterium tuberculosis*. *Biochem Biophys Res Commun* 2007; **361**: 903–909.
39 Snell EE, Shive W. Growth inhibition by analogues of pantothenic acid. Pantothenyl alcohol and related compounds. *J Biol Chem* 1945; **158**: 551–559.
40 Spry C, Chai CL, Kirk K, Saliba KJ. A class of pantothenic acid analogs inhibits *Plasmodium falciparum* pantothenate kinase and represses the proliferation of malaria parasites. *Antimicrob Agents Chemother* 2005; **49**: 4649–4657.
41 Pethe K, Swenson DL, Alonso S, Anderson J, Wang C, Russell DG. Isolation of *Mycobacterium tuberculosis* mutants defective in the arrest of phagosome maturation. *Proc Natl Acad Sci USA* 2004; **101**: 13642–13647.
42 Al-Mubarak R, Vander Heiden J, Broeckling CD, Balagon M, Brennan PJ, Vissa VD. Serum metabolomics reveals higher levels of polyunsaturated fatty acids in lepromatous leprosy: potential markers for susceptibility and pathogenesis. *PLoS Negl Trop Dis* 2011; **5**: e1303.

Supplementary Information for this article can be found on *Emerging Microbes & Infections*' website (http://www.nature.com/EMI).